Rank structures preserved by the QR-algorithm: the singular case

Steven Delvaux *Marc Van Barel*

Report TW400, August 2004

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Rank structures preserved by the QR-algorithm: the singular case

Steven Delvaux *Marc Van Barel*

Report TW400, August 2004

Department of Computer Science, K.U.Leuven

Abstract

In an earlier paper we introduced the classes of polynomial and rank structures, both of them preserved by applying a (shifted) QR-step on a matrix A. In the present paper we will further investigate the case of rank structures. We will show that even if A is a singular matrix, a new QR-iterate can be constructed having the same rank structure as the matrix A itself. To this end we will introduce the concepts of effectively eliminating QR-decompositions and sparse Givens patterns, both of them being concepts of independent interest.

Keywords : (shifted) QR-algorithm, rank structure, Givens transformations, effectively eliminating QR-decompositions, sparse Givens patterns
AMS(MOS) Classification : Primary : 15A18, Secondary : 15A23, 65F15.

Rank structures preserved by the QR-algorithm: the singular case

Steven Delvaux [*], Marc Van Barel [*]

4th August 2004

Abstract

In an earlier paper we introduced the classes of polynomial and rank structures, both of them preserved by applying a (shifted) QR-step on a matrix A. In the present paper we will further investigate the case of rank structures. We will show that even if A is a singular matrix, a new QR-iterate can be constructed having the same rank structure as the matrix A itself. To this end we will introduce the concepts of effectively eliminating QR-decompositions and sparse Givens patterns, both of them being concepts of independent interest.

Keywords: (shifted) QR-algorithm, rank structure, Givens transformations, effectively eliminating QR-decompositions, sparse Givens patterns

1 Introduction

In [3], we introduced the classes of polynomial and rank structures, both of them preserved by the (shifted) QR-algorithm. For polynomial structures, we had no problem for proving the preservation of structure. For rank structures, however, we had to make the assumption that the given matrix is nonsingular. It is the aim of this paper to remove this nonsingularity assumption.

---

[*] Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3001 Leuven (Heverlee), Belgium. email: {Steven.Delvaux,Marc.VanBarel} @cs.kuleuven.ac.be.

The research was partially supported by the Research Council K.U.Leuven, project OT/00/16 (SLAP: Structured Linear Algebra Package), by the Fund for Scientific Research–Flanders (Belgium), projects G.0078.01 (SMA: Structured Matrices and their Applications), G.0176.02 (ANCILA: Asymptotic aNalysis of the Convergence behavior of Iterative methods in numerical Linear Algebra), G.0184.02 (CORFU: Constructive study of Orthogonal Functions) and G.0455.0 (RHPH: Riemann-Hilbert problems, random matrices and Padé-Hermite approximation), and by the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, project IUAP V-22 (Dynamical Systems and Control: Computation, Identification & Modelling). The scientific responsibility rests with the authors.

To make the paper self-contained, we will repeat now several concepts. We start with the two defining equations of the shifted QR-algorithm. These equations show how to obtain from the matrix $A^{(\nu)} \in \mathbb{C}^{n\times n}$ a new iterate $A^{(\nu+1)}$:

$$A^{(\nu)} - \lambda I = QR \tag{1}$$

$$A^{(\nu+1)} = RQ + \lambda I, \tag{2}$$

where $\lambda \in \mathbb{C}$ is called the *shift*, Q is unitary and R upper triangular. As it is known, by appropriately choosing shifts the matrices $A^{(\nu)}$ converge to (block) upper triangular form, or to diagonal form in the Hermitian case, and hence the QR-algorithm can be used to determine the eigenvalues of a given matrix $A = A^{(0)}$.

We recall the definition of rank structure introduced in [3].

Definition 1 *We define a rank structure on $\mathbb{C}^{n\times n}$ as a collection of so-called structure blocks $\mathcal{R} = \{\mathcal{B}_k\}_k$. Each structure block $\mathcal{B}_k$ is characterized as a 4-tuple*

$$\mathcal{B}_k = (i_k, j_k, r_k, \lambda_k),$$

where i_k is the row index, j_k the column index, r_k the rank upper bound and $\lambda_k \in \mathbb{C}$ is called the shift element of $\mathcal{B}_k$. We say a matrix $A \in \mathbb{C}^{n\times n}$ to satisfy the rank structure if for each k,

$$\operatorname{Rank} A_k(i_k : n, 1 : j_k) \leq r_k, \quad \textit{where } A_k = A - \lambda_k I.$$

Given some rank structure $\mathcal{R}$, we denote by $\mathcal{M}_{\mathcal{R}}$, or shortly $\mathcal{M}$ the set of matrices in $\mathbb{C}^{n\times n}$ which satisfy the structure. As a special case, when all structure blocks $\mathcal{B}_k$ have shift λ_k equal to zero, then we speak about a pure rank structure on $\mathbb{C}^{n\times n}$. We denote such a structure by $\mathcal{R}_{\text{pure}}$, and we use the notation $\mathcal{M}_{\mathcal{R}_{\text{pure}}}$, or shortly $\mathcal{M}_{\text{pure}}$ to denote the class of matrices satisfying it.

A classical example of rank structure is given by the class of Hessenberg matrices. Another example is the class of lower semiseparable matrices (the latter being matrices A for which every submatrix that can be taken out of the lower triangular part of A, is of rank at most 1): see Figure 1. Hessenberg and semiseparable matrices are both examples of *pure* rank structure. An example of non-pure rank structure is given by the class of lower semiseparable *plus diagonal* matrices; here the diagonal correction $\Lambda = \operatorname{diag}(\lambda_i)_{i=1}^{n}$ is an integral part of the structure. For further examples, we refer to the references [6, 1, 4, 2, 8, 7, 5].

Every non-pure structure block $\mathcal{B}_k$ which intersects the diagonal, induces several pure structure blocks lying below the diagonal: see Figure 2. In particular, setting $l_k = \min\{j_k, i_k - 1\}$, then we define the *induced left pure structure block* of $\mathcal{B}_k$ as the 4-tuple $\mathcal{B}_{\text{left},k} = (i_k, l_k, r_k, 0)$ if $\lambda_k \neq 0$. Correspondingly, the subset of $\mathcal{I} := \{1, \ldots, n\}$ consisting of the indices $\{1, \ldots, l_k\}$ will be denoted by $\mathcal{I}_{\text{left},k}$. For a *pure* structure block $\mathcal{B}_k$ (i.e. $\lambda_k = 0$), we adapt the definition by setting $l_k := j_k$, and thus in particular $\mathcal{B}_{\text{left},k} = (i_k, j_k, r_k, 0)$, which is the structure block $\mathcal{B}_k$ itself.

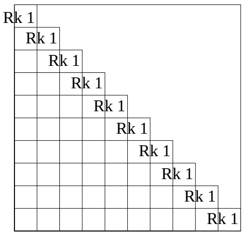


Figure 1: The figure shows the rank structure $\mathcal{R}_{\text{pure}} = \{\mathcal{B}_k\}_{k=1}^n$ which defines the class of lower semiseparable matrices. The notation Rk 1 is used to denote 'rank at most 1'.

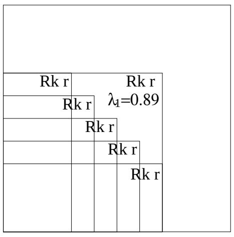


Figure 2: Example of induced pure structure: the huge structure block $\mathcal{B}_1$ with shift $\lambda_1 = 0.89$ induces 5 pure structure blocks just below the diagonal, following a kind of staircase form. The leftmost of them is called the induced left pure structure block of $\mathcal{B}_1$.

We may note that our structure blocks can be seen as an intrinsic generalization of the 'shifted' QR-algorithm, in the sense that every block is allowed to have its own shift element. As a consequence, to investigate the preservation of rank structures, it follows obviously from the QR-equations (1) and (2) that we are allowed to forget about the shift λ which is built in in the QR-algorithm, and instead just absorb it into the structure. Hence in the sequel, all our theorems will be stated for the QR-algorithm without shift.

Given a matrix A, then we define $\mathcal{I}_{\text{dep},A}$ to be the subset of $\mathcal{I} := \{1, \ldots, n\}$ consisting of the indices of all columns of A which can be written as a linear combination of the previous columns. Then we recall the following theorem from [3].

Theorem 2 *Let $\mathcal{B}_k$ be a structure block and let $A \in \mathcal{M}_{\mathcal{B}_k}$ be an arbitrary matrix, possibly singular. Then by applying a QR-step without shift on A, the rank upper bound r_k of $\mathcal{B}_k$ can increase by at most $\#(\mathcal{I}_{\text{dep},A} \cap \mathcal{I}_{\text{left},k})$.*

In particular, this theorem implies that for a nonsingular matrix A, the rank structure of A is always preserved by the QR-algorithm. On the other hand, for a *singular* matrix A the preservation of structure is not guaranteed.

In this paper we will show how to avoid the bottleneck of Theorem 2 in case A is a singular matrix. For this, we must remark that Theorem 2 works for an arbitrary new QR-iterate of A, of which there can be a lot since a singular matrix A may have many QR-factorizations $A = QR$. The point will be to choose a suitable QR-factorization, preserving structure. In this paper, we will introduce two concepts for obtaining such a QR-decomposition: effectively eliminating QR-decompositions and sparse Givens patterns. These tools will not only be interesting from a computational point of view (in terms of numerical efficiency), but also from a theoretical point of view, as will become clear throughout the paper.

2 Givens transformations

In this section, we will introduce some preliminary results about Givens transformations. Given a matrix A, we will search a QR-decomposition by solving

$$\left\{ \begin{array}{lcl} Q^H A & = & R \\ Q^H & = & (G^{(n-1)}_{n-1,n}) \ldots (G^{(2)}_{2,3} \ldots G^{(2)}_{n-1,n})(G^{(1)}_{1,2} \ldots G^{(1)}_{n-1,n}), \end{array} \right. \tag{3}$$

where each $G^{(j)}_{i-1,i}$ is an (embedded) Givens transformation acting on rows $i-1, i$ of the matrix A. We may note that (3) is not a restriction, since *any* unitary matrix Q^H can be factored in this way: we will come back to this later (see Corollary 15).

Consider for example the rightmost factor $G^{(1)}_{n-1,n}$ of (3): this Givens transformation is acting on rows $n-1$ and n, and it will serve to create a zero on the $(n, 1)$ position of the matrix A. Proceeding in this way, the factors $G^{(j)}_{i-1,i}$ will subsequently create zeros in the matrix A, hereby transforming it into an upper triangular matrix $Q^H A = R$. The $G^{(j)}_{i-1,i}$ with fixed j will be said to be in the 'jth step' of this process; they will have to make the jth column upper triangular.

Let us treat this in a formal way. Suppose that when solving (3), we applied on A all the Givens transformations $G^{(\tilde{j})}_{\tilde{i}-1,\tilde{i}}$ with either $\tilde{j} < j$, or $\tilde{j} = j$ and $\tilde{i} > i$. Then we must find $G^{(j)}_{i-1,i}$. Denoting with $\begin{bmatrix} a \\ b \end{bmatrix}$ the 2-vector containing the (updated) $(i-1, j)$ and (i, j) elements of A, we claim that $G^{(j)}_{i-1,i}$ (in its non-embedded, i.e. 2×2 form) must be chosen such that

$$G^{(j)}_{i-1,i} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} s \\ 0 \end{bmatrix}, \tag{4}$$

with $|s|^2 = |a|^2 + |b|^2$. Indeed: if this would not be the case, then clearly the $G^{(j)}_{\tilde{i}-1,\tilde{i}}$, $\tilde{i} = n, \ldots, j+1$ would leave the jth column of A in non-upper triangular form, thus with the (updated) vector $A(j+1 : n, j)$ having 2-norm different from zero. But then this 2-norm can not be changed anymore by the $G^{(\tilde{j})}_{\tilde{i}-1,\tilde{i}}$

with $\tilde{j} > j$, since these Givens transformations act only on rows $j+1, \ldots, n$, and thus we would end up with a matrix $Q^H A = R$ which is not upper triangular, yielding a contradiction. This proves (4).

Defining

$$H := \frac{1}{\sqrt{|t|^2+1}} \begin{bmatrix} 1 & \bar{t} \\ -t & 1 \end{bmatrix}, \tag{5}$$

with $t := b/a$, then the general solution of (4) (assuming $(a,b) \neq (0,0)$) is given by

$$G^{(j)}_{i-1,i} = \operatorname{diag}(\epsilon_1, \epsilon_2) H, \tag{6}$$

with $|\epsilon_1| = |\epsilon_2| = 1$. In particular, $G^{(j)}_{i-1,i}$ is essentially unique, i.e. unique up to the multiplication with the unitary diagonal matrix $\operatorname{diag}(\epsilon_1, \epsilon_2)$ in (6).

There is one exception to this essential uniqueness, namely when both a and b are zero (we refer to this case as a $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$-situation). In this case we can make any random choice for the unitary matrix $G^{(j)}_{i-1,i} \in \mathbb{C}^{2\times 2}$ to solve (4).

Let us resume the above discussion.

Definition 3 *We denote by* $A^{(1:j)}$ *the intermediate matrix which is obtained from* A *'after* j *steps', i.e. by applying all the* $G^{(\tilde{j})}_{\tilde{i}-1,\tilde{i}}$ *with* $\tilde{j} \leq j$ *in (3).*

Lemma 4

1. *The Givens transformations* $G^{(j)}_{i-1,i}$ *solve (3) if and only if each* $G^{(j)}_{i-1,i}$ *eliminates the* (i,j) *element of the matrix* A*. In particular,* $G^{(j)}_{i-1,i}$ *will be essentially unique except when we meet a* $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$*-situation.*
2. *We can choose* $G^{(j)}_{i-1,i} = I_2$ *for certain* i *and* j *if and only if the elements on positions* $(i,j), \ldots, (n,j)$ *of* $A^{(1:j-1)}$ *are equal to zero.*

PROOF.

1. This follows from the above discussion.
2. Suppose that $G^{(j)}_{i-1,i} = I_2$ for certain i and j. Then it follows that in the jth step, the transformations $G^{(j)}_{\tilde{i}-1,\tilde{i}}$ can not change the 2-norm of the subcolumn $A^{(1:j-1)}(i:n, j)$. But by part 1, we know that there have to be created zeros on these positions; hence our assumption $G^{(j)}_{i-1,i} = I_2$ can only hold if $A^{(1:j-1)}(i:n,j)$ consisted entirely of zeros. Conversely, if these zeros were standing there, then by part 1 we can clearly choose $G^{(j)}_{\tilde{i}-1,\tilde{i}} = I_2$ for $\tilde{i} = n, \ldots, i$. □

3 Effectively eliminating QR-decompositions

We already observed that in the singular case, one has to search for a suitable, non-random QR-decomposition $A = QR$ in order to preserve a given structure block $\mathcal{B}_k$. In this section we will introduce a first tool for solving this problem: the concept of effectively eliminating QR-decompositions.

The idea of the following definition is to obtain a solution of (3) which is as efficient as possible, i.e. to require that each non-trivial $G^{(j)}_{i-1,i}$ should really contribute in the process of making A upper triangular.

Definition 5 *We say a QR-decomposition $A = QR$ to be effectively eliminating if for a certain unitary diagonal matrix D, $D^H Q^H$ can be written as in (3), but this time satisfying the constraint: for each (i, j), either $G^{(j)}_{i-1,i} = I_2$, or there exists an index $l \geq j$ such that, just before applying $G^{(j)}_{i-1,i}$, the (updated) matrix $A(i-1 : n, 1 : l)$ consisted entirely of zeros except for its (i, l) and (possibly) $(i-1, l)$ elements (the former lying in the strictly lower triangular part of A), and such that $G^{(j)}_{i-1,i}$ is chosen to eliminate the (i, l) element.*

Thus each Givens transformation $G^{(j)}_{i-1,i}$ should be chosen to be either I_2, or to realize a transition

$$\begin{bmatrix} 0 & \dots & 0 & a \\ 0 & \dots & 0 & b \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & \dots & 0 & s \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix}, \qquad (7)$$

where a, b are the (updated) $(i-1, l)$ and (i, l) elements, where $l \geq j$ and with $b \neq 0$ lying in the strictly lower triangular part of A.

Remark 6

1. *The index l is certainly unique, but it does not always exist: for example it does never exist when A is an upper triangular matrix, or when we meet a situation where $b = 0$ and $a \neq 0$. Note that in the case where l does not exist, Definition 5 necessarily leads to choosing $G^{(j)}_{i-1,i} = I_2$.*

2. *An alternative for Definition 5 would be to replace the condition '$G^{(j)}_{i-1,i} = I_2$' by the condition that $G^{(j)}_{i-1,i}$ is a 2 by 2 unitary diagonal matrix* $\operatorname{diag}(\epsilon_1, \epsilon_2)$*. This would have the advantage that we could forget about the unitary diagonal matrix D occurring in Definition 5. Nevertheless, we preferred to use the condition $G^{(j)}_{i-1,i} = I_2$ since it is more 'expressive': the factor $G^{(j)}_{i-1,i}$ can then really be 'skipped' in (3).*

3. *Effectively eliminating QR-decompositions may appear in several variants which we call lazy, preparative and intermediate. We will illustrate this in the next example.*

As an example, we define

$$A = \begin{bmatrix} 0 & \times & \times \\ 0 & 1 & \times \\ 0 & 1 & \times \end{bmatrix}, \tag{8}$$

where the $\times \in \mathbb{C}$ are arbitrary. We can then exploit the freedom when meeting a $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$-situation to solve (3) as

$$\begin{aligned} & G_{2,3}^{(1)} = G_{1,2}^{(1)} = I_2, \quad G_{2,3}^{(2)} = H \quad \text{(lazy variant)} \\ \text{or} \quad & G_{2,3}^{(1)} = H, \quad G_{1,2}^{(1)} = G_{2,3}^{(2)} = I_2 \quad \text{(preparative variant)}, \end{aligned}$$

where $H := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. Thus the lazy variant chooses $G_{i-1,i}^{(j)} = I_2$ ('skipping' the Givens transformation) as soon as possible, while the preparative variant only chooses $G_{i-1,i}^{(j)} = I_2$ if the index l of Definition 5 does not exist; in this case we had $l = 2$. We may mention that for larger examples, also intermediate variants exist between lazy and preparative.

It should be noted that in the above example, the lazy and preparative variant are in fact just two different ways of factorizing a *single* QR-decomposition $A = QR$ as a Givens product (3). We will prove a more general version of this result.

Theorem 7 *The effectively eliminating QR-decomposition of A is essentially unique, i.e. for two effectively eliminating QR-decompositions $A = Q_1 R_1$ and $A = Q_2 R_2$, we have $Q_1 = Q_2 D$ for a certain unitary diagonal matrix D.*

PROOF. Given an effectively eliminating QR-decomposition $A = QR$, we will first show that it can be refactored to be in the *preparative variant.* Thus suppose that at a certain point, we chose $G_{i-1,i}^{(j)} = I_2$ (being lazy) although an index l existed which satisfies the constraints in Definition 5. Then by the existence of this index l, it follows from Definition 5 that the next Givens transformations $G_{\tilde{i}-1,\tilde{i}}^{(\tilde{j})}$ were not allowed to involve the $(i-1)$th or ith row, until for certain $\tilde{j} > j$, $G_{i-1,i}^{(\tilde{j})}$ was chosen to eliminate the (i,l) element. But then by the commuting of two Givens transformations acting on strictly disjoint row indices, we can just redefine $G_{i-1,i}^{(j)} := G_{i-1,i}^{(\tilde{j})} G_{i-1,i}^{(j)}$ and $G_{i-1,i}^{(\tilde{j})} := I_2$. Repeating this argument several times, the reduction to the preparative variant can be realized.

Next, we will show that the effectively eliminating QR-decomposition in the preparative variant is essentially unique. Since the preparative variant tells us exactly what to do when meeting a $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$-situation, the only freedom that we have are the factors $\mathrm{diag}(\epsilon_1, \epsilon_2)$ in each solution (6). We claim that their influence is limited to the multiplication of Q^H with a unitary diagonal matrix

D^H on the left. Indeed: suppose by induction that we are choosing some non-trivial $G^{(j)}_{i-1,i}$, and that the $(i-1)$th row has been multiplied by ϵ_a, and the ith row by ϵ_b, for certain $|\epsilon_a| = |\epsilon_b| = 1$. Let us define the ratio $t := b/a$ which would occur in (5) in case that $\epsilon_a = \epsilon_b = 1$. The actual value of this ratio must then be corrected by the factor ϵ_b/ϵ_a; this can be realized by decomposing the matrix H of (5) as

$$H = \operatorname{diag}(\epsilon_a, \epsilon_b) \frac{1}{\sqrt{|t|^2+1}} \begin{bmatrix} 1 & \overline{t} \\ -t & 1 \end{bmatrix} \operatorname{diag}(\overline{\epsilon}_a, \overline{\epsilon}_b). \tag{9}$$

Thus the application of H can be decomposed into three steps: the first step is to apply $\operatorname{diag}(\overline{\epsilon}_a, \overline{\epsilon}_b)$, hence annihilating the influence of ϵ_a, ϵ_b. Next, the central factor of (9) is precisely the matrix H corresponding to the case $\epsilon_a = \epsilon_b = 1$. Finally, we apply $\operatorname{diag}(\epsilon_a, \epsilon_b)$, which can be absorbed into the factor $\operatorname{diag}(\epsilon_1, \epsilon_2)$ of (6) to yield *again* a scaling of the rows. Repeating this argument for all non-trivial $G^{(j)}_{i-1,i}$, the essential uniqueness is proved. $\square$

4 Sparse Givens patterns

In this section we will introduce a second tool for obtaining a 'suitable' QR-decomposition of a matrix. The setting is a bit different than in the previous section, in the sense that we will work now with a pure structure $\mathcal{R}_{\text{pure}}$ rather than a single matrix A.

Given a pure structure block $\mathcal{B}_k$, we will define two index sets.

Definition 8 *We define $\mathcal{I}^2 = \{(i,j) \in \mathcal{I} \times \mathcal{I} \mid i > j\}$, i.e. precisely the set of indices for which $G^{(j)}_{i-1,i}$ is a Givens transformation occurring in (3). Then if $\mathcal{B}_k = (i_k, j_k, r_k)$ is a pure structure block, we define the two index sets*

$$\mathcal{I}^2_{\text{Prepare},k} = \{(i,j)|\ 1 \leq j \leq r_k \text{ and } i_k + j \leq i \leq n\} \subseteq \mathcal{I}^2 \qquad \textit{(staircase)},$$
$$\mathcal{I}^2_{\text{Skip},k} = \{(i,j)|\ r_k + 1 \leq j \leq j_k \text{ and } i_k + r_k \leq i \leq n\} \subseteq \mathcal{I}^2 \qquad \textit{(rectangle)}.$$

Note that the rectangular set $\mathcal{I}^2_{\text{Skip},k}$ is obtained from $\mathcal{B}_k$ by just dropping the first r_k rows and columns: see Figure 3.

Remark 9 *For the above definition to be completely meaningful, we should have that $\{\mathcal{I}^2_{\text{Prepare},k}, \mathcal{I}^2_{\text{Skip},k}\} \subseteq \mathcal{I}^2$ for all k. It can be checked graphically that this is satisfied if and only if the top right element of each set $\mathcal{I}^2_{\text{Skip},k}$ belongs to $\mathcal{I}^2$, thus if*

$$i_k + r_k > j_k \tag{10}$$

for all k. In fact, this condition is equivalent with the more 'intrinsic' statement that the structure $\mathcal{R}_{\text{pure}}$ does not imply singularity, i.e. that the structure $\mathcal{R}_{\text{pure}}$ itself does not force the matrix to be singular. From now on, we will always suppose this condition to hold. If this is not the case, then the next theorems

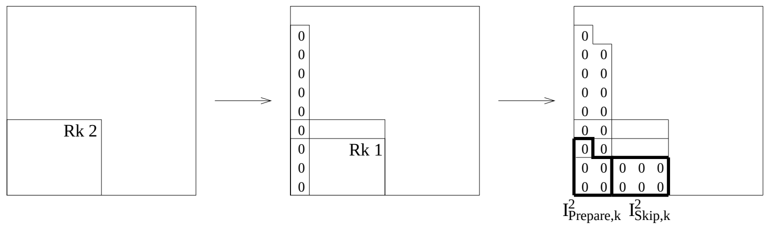


Figure 3: Given the Rk 2 structure block $\mathcal{B}_k$ in the left picture. After applying the $G^{(1)}_{i-1,i}$ (suitably chosen), zeros are introduced in the first column and the structure block $\mathcal{B}_k$ is transformed into a new structure block with one row less and rank one less, i.e. rank 1. After applying the $G^{(2)}_{i-1,i}$, a new structure block is obtained with again a row less, but now with rank equal to zero. Hence there has been automatically created a 2×3 block of zeros on the bottom of A. Thus the $G^{(j)}_{i-1,i}$ intended to create zeros on these positions are unnecessary and can be 'skipped', hence explaining the notation $\mathcal{I}^2_{\text{Skip},k}$.

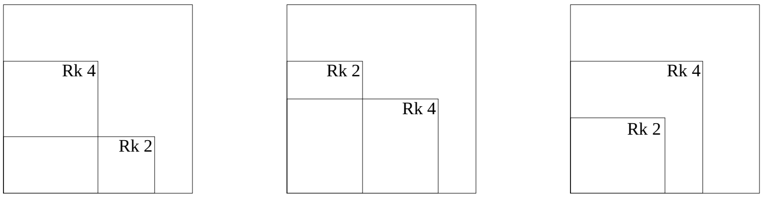
will not work: a possible remedy is then to restrict to the maximal induced structure which does not imply singularity, and then to work instead with this (slightly weaker) structure.

The idea behind Definition 8 is the following: for any $A \in \mathcal{M}_{\text{pure}}$, a QR-decomposition can be found such that $G^{(j)}_{i-1,i} = I_2$ (skipping this Givens transformation) for all $(i,j) \in \bigcup_k \mathcal{I}^2_{\text{Skip},k}$. Loosely speaking, this idea can be realized by taking an effectively eliminating QR-decomposition of A, written in an 'intermediate' variant, by choosing the preparative variant for all $(i,j) \in \mathcal{I}^2_{\text{Prepare},k}$ and the lazy variant for all $(i,j) \in \mathcal{I}^2_{\text{Skip},k}$. The reader should check the meaning of this for a block of zeros ($r_k = 0$), and for the general idea illustrated in Figure 3.

However, we should be careful about one additional element: the *interaction* between the different structure blocks $\mathcal{B}_k$, $\mathcal{B}_{\tilde{k}}$, meaning that $\mathcal{I}^2_{\text{Skip},k}$ may possibly overlap with $\mathcal{I}^2_{\text{Prepare},\tilde{k}}$ for certain $k, \tilde{k}$, which may lead to a conflict. Obviously, such overlaps are only possible for structure blocks $\mathcal{B}_k$, $\mathcal{B}_{\tilde{k}}$ having rank upper bounds satisfying $r_{\tilde{k}} > r_k$. The actual flavour of the overlap depends then on the relative position of $\mathcal{B}_k$, $\mathcal{B}_{\tilde{k}}$, as shown in Figure 4.

Now we will suppose that there is no relative position as in type c, i.e. that there are no two structure blocks which are contained in each other.

Definition 10 *Let $\mathcal{R}_{\text{pure}} = \{\mathcal{B}_k\}_k$ be a pure structure, not implying singularity and not containing two structure blocks of type c. We say a QR-decomposition $A = QR$ to satisfy the* sparse Givens pattern induced by $\mathcal{R}_{\text{pure}}$ *if for a certain unitary diagonal matrix D, $D^H Q^H$ can be written as in (3), but this time with*



Type a Type b Type c

Figure 4: The three possible types of relative position of two structure blocks $\mathcal{B}_k$ and $\mathcal{B}_{\tilde{k}}$ with $r_{\tilde{k}} > r_k$; in the picture we have $r_k = 2$ and $r_{\tilde{k}} = 4$. By convention, the case where the column indices $j_k = j_{\tilde{k}}$ is defined to be of type a (instead of type c), and the case where the row indices $i_k = i_{\tilde{k}}$ is defined to be of type b (instead of type c).

$G^{(j)}_{i-1,i} = I_2$ *for all* (i, j) *for which*

$$\exists k : (i, j) \in \mathcal{I}^2_{\mathrm{Skip},k} \setminus \bigcup_{\tilde{k}, type\ b} \mathcal{I}^2_{\mathrm{Prepare},\tilde{k}}, \tag{11}$$

where the union $\bigcup_{\tilde{k},type\ b} \mathcal{I}^2_{\mathrm{Prepare},\tilde{k}}$ *is taken over all indices* $\tilde{k}$ *satisfying* $r_{\tilde{k}} > r_k$ *and* $i_{\tilde{k}} \geq i_k$ *(type b).*

Remark 11 *If there is no relative position as in type b, then condition (11) simplifies to skipping all Givens transformations* $G^{(j)}_{i-1,i}$ *with*

$$(i, j) \in \bigcup_{k} \mathcal{I}^2_{\mathrm{Skip},k}. \tag{12}$$

In particular, the simplification to (12) holds for the case of a single structure block (as in Figure 3), or more generally when all structure blocks $\mathcal{B}_k$ *have the same rank upper bound* r_k*, since we excluded this situation in all types a, b and c. Thus (12) will be valid for all kinds of Hessenberg and lower semiseparable structures.*

Now we prove

Theorem 12 *Under the same conditions as in Definition 10, we have that*

1. *(Existence:) if* $A \in \mathcal{M}_{\mathrm{pure}}$*, then each effectively eliminating QR-decomposition of* A *satisfies the sparse Givens pattern induced by* $\mathcal{R}_{\mathrm{pure}}$*;*
2. *(Complete characterization:) furthermore, we have that* $A \in \mathcal{M}_{\mathrm{pure}}$ *if and only if a QR-decomposition* $A = QR$ *exists which satisfies the sparse Givens pattern induced by* $\mathcal{R}_{\mathrm{pure}}$*.*

PROOF.

1. Let $A \in \mathcal{M}_{\text{pure}}$. We claim that the following conditions yield an effectively eliminating QR-decomposition of A:

   E1 For all $(i,j) \in \bigcup_k \mathcal{I}^2_{\text{Prepare},k}$, we set

$$j_{\max} := \max_{k|(i,j)\in\mathcal{I}^2_{\text{Prepare},k}} j_k,$$

   and we choose $G^{(j)}_{i-1,i}$ to eliminate the (i,l) element, where $l \leq j_{\max}$ is the smallest column index for which not both the (updated) $(i-1,l)$ and (i,l) elements are equal to zero. If there is no such column index $l \leq j_{\max}$, or if the element (i,l) is already zero, then we are lazy and choose $G^{(j)}_{i-1,i} = I_2$.

   E2 Inside the set $\mathcal{I}^2 \setminus \bigcup_k \mathcal{I}^2_{\text{Prepare},k}$, we apply the lazy variant of effectively eliminating.

   Indeed: to establish the effectively eliminating character of E1, E2, it is sufficient to show that an index l satisfying the constraint in E1, will also satisfy the constraint in Definition 5, i.e. that the (i,l) and (possibly) $(i-1,l)$ elements are the only non-zero elements in the entire (updated) submatrix $A(i-1:n, 1:l)$. This is surely satisfied for the elements on the $(i-1)$th and ith row, by construction of E1. Then suppose by contradiction that some of the $(\tilde{i},\tilde{j})$ elements are zero, where $\tilde{i} > i$ and $\tilde{j} \leq l$. Let us choose such an element for which $\tilde{j}$ is minimal. Then we claim that $G^{(j)}_{\tilde{i}-1,\tilde{i}}$ has not been chosen correctly: this follows since surely $\tilde{j} \leq l \leq j_{\max}$, and since $j_{\max}$ is an increasing function when i increases (by the shape of the sets $\mathcal{I}^2_{\text{Prepare},k}$), condition E1 implies that $G^{(j)}_{\tilde{i}-1,\tilde{i}}$ should have been chosen to eliminate this non-zero $(\tilde{i},\tilde{j})$ element, yielding a contradiction.

   Thus we see that E1, E2 are effectively eliminating; hence by the essential uniqueness of Theorem 7, it is sufficient to prove the theorem for these conditions E1, E2. Now let us fix a structure block $\mathcal{B}_k$. Let $1 \leq j \leq r_k$ and suppose that we applied on A the first $j-1$ steps of (3), leading to the updated matrix

$$A^{(1:j-1)} = \begin{bmatrix} R & X & X \\ 0 & X & X \\ 0 & S & X \end{bmatrix}$$

   with R upper triangular of size $j-1$ by $j-1$, and with the submatrix S having row indices $i_k+j-1,\ldots,n$ and column indices $j,\ldots,j_k$ (thus $j-1$ rows and columns less than $\mathcal{B}_k$). Suppose that S is of rank at most r, for certain $r > 0$. Then by applying the factor $G := G^{(j)}_{i_k+j-1,i_k+j} \cdots G^{(j)}_{n-1,n}$, we get

$$GS = \left[\begin{array}{c|ccc} s & \times & \ldots & \times \\ \hline & & & \\ 0 & \times & \ldots & \times \\ \vdots & \vdots & & \vdots \\ 0 & \times & \ldots & \times \end{array}\right] = \mathrm{Rk}\ r, \tag{13}$$

which is still of rank at most r, and where $|s|$ is the 2-norm of the column $S(:,1)$. We supposed here that $s \neq 0$; if not, then supposing that S has $c > 0$ zero columns on the left, condition E1 will lead to exactly the same situation as in (13), except for c additional zero columns added on the left. Then it is easy to check that the Rk r condition and the fact that $s \neq 0$ imply the bottom submatrix of GS in (13) (obtained by dropping the first row) to be of the form Rk $(r-1)$.

By subsequently using this observation, it follows that during the process of applying E1, E2 to make A upper triangular, the original structure blocks $\mathcal{B}_k$ transform into new structure blocks with one row less, and with rank diminished by one (this process was already illustrated in Figure 3). After at most r_k steps, this yields a rectangular block of zeros on the bottom of $A^{(1:r_k)}$, with coordinates being precisely the set $\mathcal{I}^2_{\mathrm{Skip},k}$.

Having established this, we will now be able to show that E1, E2 satisfy the sparse Givens pattern. Thus we must show that $G^{(j)}_{i-1,i} = I_2$ for all (i,j) for which for certain k, $(i,j) \in \mathcal{I}^2_{\mathrm{Skip},k}$ and $(i,j) \notin \mathcal{I}^2_{\mathrm{Prepare},\tilde{k}}$ for all $\tilde{k}$ satisfying type b. These assumptions imply that (i,j) can only belong to sets $\mathcal{I}^2_{\mathrm{Prepare},\tilde{k}}$ with $k, \tilde{k}$ related as in type a. Hence by definition of type a, we have that

$$j_{\max} := \max_{\tilde{k} \mid (i,j) \in \mathcal{I}^2_{\mathrm{Prepare},\tilde{k}}} j_{\tilde{k}} \leq j_k.$$

But then the presence of the zero block corresponding to $\mathcal{I}^2_{\mathrm{Skip},k}$ implies that for every column index $l \leq j_{\max} \leq j_k$, both the $(i-1,l)$ and (i,l) elements are already zero. Hence the lazy character of conditions E1, E2 implies us to choose $G^{(j)}_{i-1,i} = I_2$, thus establishing the sparse Givens pattern.

2. The implication ⇒ (Existence) follows directly from part 1. To prove the implication ⇐ (Complete characterization), let us suppose that $A = QR$ satisfies the sparse Givens pattern induced by $\mathcal{R}_{\mathrm{pure}}$. Let $\mathcal{B}_k \in \mathcal{R}_{\mathrm{pure}}$ be a structure block. Let us consider the *top tow* of the rectangle $\mathcal{I}^2_{\mathrm{Skip},k}$, i.e.

   $$\{(i,j) \mid i = i_k + r_k \text{ and } j \geq r_k + 1\}.$$

   Then we claim that (*): the indices of this top row can not belong to any of the sets $\mathcal{I}^2_{\mathrm{Prepare},\tilde{k}}$, where $\mathcal{B}_{\tilde{k}}$ is a structure block as in type b. Assuming this for the moment, the sparse Givens pattern implies us to choose $G^{(j)}_{i-1,i} = I_2$ in the complete top row of $\mathcal{I}^2_{\mathrm{Skip},k}$. Hence from Lemma

4.2 it follows that after r_k steps, a rectangular block of zeros must have been present on the bottom of $A^{(1:r_k)}$, with indices given by $\mathcal{I}^2_{\mathrm{Skip},k}$. But then clearly the presence of this zero block implies that the original matrix A must have satisfied the original structure block $\mathcal{B}_k$: this can be seen since in every application of (13), the rank of the bottom submatrix of S can decrease by *at most 1*.

Thus we would be finished if we can prove our claim (*). The reader may check this either graphically, or algebraically since if (*) is not satisfied, we would get the contradiction

$$i_k + r_k = i \geq i_{\tilde{k}} + j \geq i_{\tilde{k}} + r_k + 1 \geq i_k + r_k + 1, \tag{14}$$

where we used subsequently the assumption $i = i_k + r_k$ (top row of $\mathcal{I}^2_{\mathrm{Skip},k}$), $i \geq i_{\tilde{k}} + j$ (definition of $\mathcal{I}^2_{\mathrm{Prepare},\tilde{k}}$), the assumption that $j \geq r_k + 1$ (definition of $\mathcal{I}^2_{\mathrm{Skip},k}$) and the fact that $i_{\tilde{k}} \geq i_k$ (definition of type b). $\square$

Let us consider now the case where also type c structure blocks $\mathcal{B}_k, \mathcal{B}_{\tilde{k}}$ are allowed in the structure. We introduce the intermediate structure blocks $\mathcal{B}_{\mathrm{inter},a} := (i_{\tilde{k}}, j_k, r_{\mathrm{inter},a})$ and $\mathcal{B}_{\mathrm{inter},b} := (i_k, j_{\tilde{k}}, r_{\mathrm{inter},b})$, where $r_{\mathrm{inter},a}$ and $r_{\mathrm{inter},b}$ are rank upper bounds for $A|_{\mathcal{B}_{\mathrm{inter},a}}$ and $A|_{\mathcal{B}_{\mathrm{inter},b}}$; the value $r_{\mathrm{inter},a}$ is assumed to be the *exact* rank. These ranks must be related, as illustrated in Figure 5.

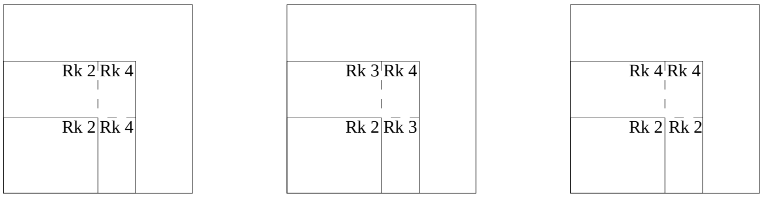


Figure 5: Given two structure blocks $\mathcal{B}_k$ and $\mathcal{B}_{\tilde{k}}$ with $r_k = 2$ and $r_{\tilde{k}} = 4$; for simplicity we assume r_k to be *exactly* 2. Then there are three possible distributions for the ranks of the two intermediate structure blocks $\mathcal{B}_{\mathrm{inter},a}$ and $\mathcal{B}_{\mathrm{inter},b}$. In the figure, the values of $r_{\mathrm{inter},a}$ in top left position are assumed to be *exact*.

We consider the problem of defining the sparse Givens pattern induced by $\mathcal{R}_{\mathrm{pure}}$, not implying singularity. Let us give a sketch of our solution to this problem. First we define the structure

$$\tilde{\mathcal{R}}_{\mathrm{pure}} = \mathcal{R}_{\mathrm{pure}} \cup \bigcup_{k,\tilde{k},\mathrm{type\ c}} \mathcal{B}_{\mathrm{inter},a},$$

with index running over all type c structure blocks. To be honest, it could be that one of the blocks $\mathcal{B}_{\mathrm{inter},a}$ implies singularity, but then we can just increase its row index until (10) is valid again; the original structure block $\mathcal{B}_{\tilde{k}}$ can then be removed since it will be automatically induced.

It is then sufficient to define the sparse Givens pattern induced by this (stronger) structure $\tilde{\mathcal{R}}_{\text{pure}}$. We do this as follows: for each index $\tilde{k}$, we *redefine* the set $\mathcal{I}^2_{\text{Prepare},\tilde{k}}$ to be $\mathcal{I}^2_{\text{Prepare},\tilde{k}} = \{(i,j)|\ r_{\text{inter},a}+1 \leq j \leq r_k \text{ and } i_k+j \leq i \leq n\} \subseteq \mathcal{I}^2$, hence obtained from the 'usual' set $\mathcal{I}^2_{\text{Prepare},\tilde{k}}$ by dropping the first $r_{\text{inter},a}$ columns, where $r_{\text{inter},a}$ is maximal over all structure blocks $\mathcal{B}_k$ for which $\mathcal{B}_k, \mathcal{B}_{\tilde{k}}$ are as in type c. This can be justified as follows: during the preparation of $\mathcal{B}_{\text{inter},a}$, *automatically* the structure block $\mathcal{B}_{\tilde{k}}$ will be prepared too. Another change is that in (11), the union should be taken over all structure blocks of types b and c.

Treating these ideas in a formal way, the definition of sparse Givens pattern induced by $\tilde{\mathcal{R}}_{\text{pure}}$ can be given in such a way that Theorem 12 remains valid. But we prefer not to reserve space for the technical details of checking this, since the transition from $\mathcal{R}_{\text{pure}}$ to $\tilde{\mathcal{R}}_{\text{pure}}$ and the appearance of all the *exact* ranks $r_{\text{inter},a}$ lead to a rather unnatural solution to this problem. Furthermore, in practice most structures will be of Hessenberg or lower semiseparable related types, hence not including type c structure blocks at all.

5 Preservation of rank structure

In this section we will prove that the concepts of effectively eliminating QR-decompositions and sparse Givens patterns are exactly the right tools for obtaining a new QR-iterate of a (possibly) singular matrix A, such that the rank structure of A is preserved. This is interesting since it shows that these concepts are not only interesting from a computational point of view (in terms of numerical efficiency), but also from a theoretical point of view.

First we will prove an auxiliary theorem, being of independent interest. We recall the following result from [3].

Theorem 13 *Let A be a nonsingular matrix, satisfying a certain structure $\mathcal{R}$. Let $\mathcal{R}_{\text{pure}}$ be the induced pure structure of $\mathcal{R}$. Then for each QR-decomposition $A = QR$, we have $Q \in \mathcal{M}_{\text{pure}}$.*

Furthermore, we will also be interested in families of QR-decompositions $A_\epsilon = Q_\epsilon R_\epsilon$, $\epsilon \in \mathbb{C}$, satisfying the conditions

- F1 A_ϵ is nonsingular, except for a finite number of $\epsilon \in \mathbb{C}$;
- F2 $A_\epsilon \in \mathcal{M}_{\text{pure}}$ for all ϵ;
- F3 $\lim_{\epsilon\to 0} Q_\epsilon = Q$, $\lim_{\epsilon\to 0} R_\epsilon = R$ and (hence) $\lim_{\epsilon\to 0} A_\epsilon = A$,

In particular, we will refer to the last condition F3 by saying that $A = QR$ is the *limit* of the family of QR-decompositions $A_\epsilon = Q_\epsilon R_\epsilon$.

Theorem 14 *Let $\mathcal{R}_{\text{pure}}$ be a pure structure which does not imply singularity. Let $A \in \mathcal{M}_{\text{pure}}$ be arbitrary, possibly singular. Then the following statements are equivalent: solving $A = QR$*

(i) satisfying the sparse Givens pattern induced by $\mathcal{R}_{\text{pure}}$;

(ii) as the limit of a family $A_\epsilon = QR_\epsilon$, satisfying the above conditions F1, F2 and F3;

(iii) as the limit of a family $A_\epsilon = Q_\epsilon R_\epsilon$, satisfying F1, F2 and F3;

(iv) with $Q \in \mathcal{M}_{\text{pure}}$.

PROOF.

a. To prove that (i) $\Rightarrow$ (ii), we give an explicit construction of such a family. We do this by defining R_ϵ from R by replacing every zero diagonal element by the parameter ϵ, and putting $A_\epsilon := QR_\epsilon$. Then obviously A_ϵ is nonsingular for all $\epsilon \neq 0$. Since A_ϵ has a QR-decomposition with the same Q-factor as in A itself, and thus satisfying the sparse Givens pattern induced by $\mathcal{R}_{\text{pure}}$, it follows from Theorem 12 that $A_\epsilon \in \mathcal{M}_{\text{pure}}$ for all $\epsilon \in \mathbb{C}$. The fact that $\lim_{\epsilon\to 0} R_\epsilon = R$ and (hence) $\lim_{\epsilon\to 0} A_\epsilon = A$ is obvious.

b. The implication (ii) $\Rightarrow$ (iii) is logically trivial.

c. To prove that (iii) $\Rightarrow$ (iv), let $A_\epsilon = Q_\epsilon R_\epsilon$ be a family satisfying F1, F2 and F3. Let $S \subseteq \mathbb{C}$ be the finite set of ϵ for which A_ϵ is singular. Fixing a certain $\epsilon \in \mathbb{C} \setminus S$, it follows from Theorem 13 that for the QR-factorization $A_\epsilon = Q_\epsilon R_\epsilon$, we must have that $Q_\epsilon \in \mathcal{M}_{\text{pure}}$. Obviously, this will then also hold for the limiting matrix $Q = \lim_{\epsilon\to 0} Q_\epsilon$.

d. Finally we prove the implication (iv) $\Rightarrow$ (i). Thus suppose that $Q \in \mathcal{M}_{\text{pure}}$ satisfies the same structure as A itself. By Theorem 12, there must be a QR-factorization $\tilde{Q}^H Q = R$ with $\tilde{Q}^H$ satisfying the sparse Givens pattern induced by $\mathcal{R}_{\text{pure}}$. Moreover, by appropriate choice of the diagonal matrix D of Definition 10, we can make the diagonal elements of R to be non-negative. But then $\tilde{Q}^H Q = R$ must be a unitary, upper triangular matrix with non-negative diagonal elements, and hence this must be the identity matrix. It follows that $Q^H = \tilde{Q}^H$, satisfying the sparse Givens pattern. $\square$

Note that for part d of the proof, we needed that the definition of sparse Givens pattern only depends on the structure $\mathcal{R}_{\text{pure}}$, and not on the particular matrix $A \in \mathcal{M}_{\text{pure}}$. This is because we were dealing here with two *different* matrices A and Q, both belonging to $\mathcal{M}_{\text{pure}}$.

We have the following corollary.

Corollary 15 *By applying Theorem 14 (iv) $\Rightarrow$ (i) to the empty structure $\mathcal{R}_{\text{pure}} = \emptyset$, we get that for every QR-decomposition $A = QR$, there exists a unitary diagonal matrix D such that $D^H Q^H$ can be written as in (3). In fact this matrix D can be absorbed into the formula (3), by appropriately updating the factors of the form $G^{(j)}_{j,j+1}$; hence we conclude that every unitary matrix Q^H can be written as a Givens product (3), a fact which we already announced earlier.*

Instead of sparse Givens patterns, let us now turn to the case of effectively eliminating QR-decompositions. We will prove a theorem very similar to Theorem 14.

Definition 16 *For a given matrix A, we define the rank structure $\mathcal{R}_{\text{pure},A}$ to be the union of all pure structure blocks which are satisfied by A, and which do not imply singularity.*

Theorem 17 *Let A be an arbitrary matrix, possibly singular. Then the following statements are equivalent: solving $A = QR$*

- *(i) in an effectively eliminating way;*
- *(ii) as the limit of a family $A_\epsilon = QR_\epsilon$, satisfying F1, F2 and F3 (with $\mathcal{M}_{\text{pure}}$ replaced by $\mathcal{M}_{\text{pure},A}$);*
- *(iii) with $Q \in \mathcal{M}_{\text{pure},A}$.*

PROOF.

a. The implications (i) $\Rightarrow$ (ii) and (ii) $\Rightarrow$ (iii) are an immediate consequence of Theorem 14.

b. For the implication (iii) $\Rightarrow$ (i), it is enough to prove that a unitary matrix Q satisfying $Q \in \mathcal{M}_{\text{pure},A}$ must be essentially unique. Suppose by induction that we have constructed columns $1, \ldots, j-1$ of the matrix Q, and that these columns were essentially unique. We must then construct the jth column of Q. There are two possiblilies: (1) the jth column of A is independent of columns $1, \ldots, j-1$ of the matrix Q, and then since the QR-equation $A = QR$ implies

$$\text{span}\{\vec{A}_1, \ldots, \vec{A}_j\} \subseteq \text{span}\{\vec{Q}_1, \ldots, \vec{Q}_j\},$$

the jth column of Q will be fixed up to a unitary scaling factor $c \in \mathbb{C}$; (2) the jth column of A is dependent of columns $1, \ldots, j-1$ of the matrix Q. But then we can consider in $\mathcal{R}_{\text{pure},A}$ the largest pure structure block $\mathcal{B}_k$ having $j_k = j$ and not implying singularity. (By (10), this means that $j_k - i_k = r_k - 1$ for this structure block). Suppose then, by contradiction, that there are two linearly independent choices $\vec{q}_1, \vec{q}_2$ that can be made for $\vec{Q}_j$. We can then extend Q by adding the columns $\vec{Q}_j = \vec{q}_1$ and $\vec{Q}_{j+1} = \vec{q}_2$. This matrix Q will satisfy the structure block $\tilde{\mathcal{B}}_k$ which is obtained from $\mathcal{B}_k$ by adding one column ($j_k := j_k + 1$); but obviously this block $\tilde{\mathcal{B}}_k$ must imply singularity, hence contradicting the fact that the columns of a unitary matrix are linearly independent.

□

Now we are ready to handle the preservation of structure.

Theorem 18 *Let $\mathcal{R}$ be a structure which does not imply singularity, and let $\mathcal{R}_{\mathrm{pure}} \subseteq \mathcal{R}$ be the induced pure structure. Let $A \in \mathcal{M}$ be arbitrary, possibly singular. Then by applying a QR-step without shift on A, satisfying the sparse Givens pattern induced by $\mathcal{R}_{\mathrm{pure}}$, we have that*

1. *the induced pure structure $\mathcal{R}_{\mathrm{pure}}$ itself will always be preserved;*
2. *all structure blocks $\mathcal{B}_k$ with shift $\lambda_k \neq 0$, and for which the induced left pure structure block of $\mathcal{B}_k$ has its maximal allowed rank, i.e. rank equal to r_k, will be preserved;*
3. *if additionally we are working with an effectively eliminating QR-decomposition of A, then the complete structure $\mathcal{R}$ will be preserved.*

PROOF.

1. From the fact that $Q \in \mathcal{M}_{\mathrm{pure}}$ (Theorem 14 (i)⇒(iv)) it follows that also $RQ \in \mathcal{M}_{\mathrm{pure}}$ since the factor R takes linear combinations of the rows, only involving 'further' rows, and hence this factor can not destroy the pure structure blocks satisfied by Q.
2. We claim that in general,

$$\mathrm{span} A|_{\mathcal{B}_{\mathrm{pure},k}} \subseteq \mathrm{span} Q|_{\mathcal{B}_{\mathrm{pure},k}} = \mathrm{Rk}\ r_k, \tag{15}$$

with Rk r_k being a matrix of rank at most r_k. Indeed: the inclusion $\subseteq$ follows by the QR-equation $A = QR$, and the second transition is again just Theorem 14 (i)⇒(iv).

Suppose then additionally that we have a structure block $\mathcal{B}_k$ with $\lambda_k \neq 0$, and such that the rank of $A|_{\mathcal{B}_{\mathrm{left},k}}$ has its maximal allowed value, i.e. equal to r_k. Then by reasons of dimension, it follows that for $\mathcal{B}_{\mathrm{left},k}$ also the inclusion $\supseteq$ must hold in (15).

Now we define a family of upper triangular matrices R_ϵ by replacing every zero diagonal element of R, standing in $\mathcal{I}_{\mathrm{left},k}$, by the parameter ϵ. This yields us a family of matrices $A_\epsilon := QR_\epsilon$, $\epsilon \in \mathbb{C}$, satisfying the conditions

F1 $\mathcal{I}_{\mathrm{dep},A_\epsilon} \cap \mathcal{I}_{\mathrm{left},k} = \emptyset$, except for $\epsilon = 0$;
F2 $A_\epsilon \in \mathcal{M}_{\mathcal{B}_k}$ for all ϵ;
F3 $\lim_{\epsilon \to 0} R_\epsilon = R$.

Conditions F1 and F3 are obvious by construction. To prove condition F2, let us first show that

$$\mathrm{span} A|_{\mathcal{B}_{\mathrm{left},k}} = \mathrm{span} Q|_{\mathcal{B}_{\mathrm{left},k}} = \mathrm{span} A_\epsilon|_{\mathcal{B}_{\mathrm{left},k}}, \tag{16}$$

for all $\epsilon \neq 0$. Here the first equality is just (15), where we already remarked also the inclusion $\supseteq$ to hold. The second equality follows in a completely similar way, this time using the QR-equation $A_\epsilon = QR_\epsilon$, together with the nonsingularity of R_ϵ in $\mathcal{I}_{\text{left},k}$. Then since by construction the A_ϵ, $\epsilon \neq 0$ can only differ from A in the columns with index in $\mathcal{I}_{\text{left},k}$, condition F2 is just a consequence of (16). Thus we established F1, F2 and F3.

Now using F1, F2 and F3, we can easily finish the proof: let $\epsilon \in \mathbb{C} \setminus \{0\}$, then F2 induces $A_\epsilon \in \mathcal{M}_{\mathcal{B}_k}$, and thus from F1 and Theorem 2 it follows that also the new QR-iterate $R_\epsilon Q \in \mathcal{M}_{\mathcal{B}_k}$. Clearly the same must then be true for the limit $RQ = (\lim_{\epsilon\to 0} R_\epsilon)Q = \lim_{\epsilon\to 0}(R_\epsilon Q)$.

3. Theorem 17 (i) $\Rightarrow$ (iii) implies that for an effectively eliminating QR-decomposition, we have the equality Rank $Q|_{\mathcal{B}_{\text{left},k}} =$ Rank $A|_{\mathcal{B}_{\text{left},k}}$. Then by reasons of dimension, again the inclusion $\supseteq$ must hold for $\mathcal{B}_{\text{left},k}$ in (15); the rest of the proof is identical to the one of part 2. $\square$

Worked example: (usual) lower semiseparable plus diagonal matrices. The class of (usual) lower semiseparable plus diagonal matrices is defined as $\mathcal{R} = \{\mathcal{B}_k\}_{k=1}^n$ with $\mathcal{B}_k = (i_k, j_k, r_k, \lambda_k) = (k, k, 1, \lambda_k)$. The shift elements λ_k are called here *diagonal elements*, and we assume them to be fixed; here λ_1 and λ_n could be called 'pseudo'diagonal elements since their value is of no actual importance.

Obviously, the induced pure structure of $\mathcal{R}$ is given by $\mathcal{R}_{\text{pure}} = \{\mathcal{B}_k\}_{k=1}^{n-1}$ with $\mathcal{B}_k = (i_k, j_k, r_k) = (k+1, k, 1)$. Here we assumed that $\lambda_k \neq 0$, $k = 2, \ldots, n-1$; if this is not the case then some of the structure blocks of $\mathcal{R}_{\text{pure}}$ must be enlarged again to involve the diagonal.

To compute the sparse Givens pattern induced by $\mathcal{R}_{\text{pure}}$, we may note that since the rank upper bounds r_k are all the same, the definition reduces to just skipping all $G^{(j)}_{i-1,i}$ with $(i,j) \in \bigcup_k \mathcal{I}^2_{\text{Skip},k}$; see (12). It is easy to see that this implies a solution to (3) with

$$Q^H = (G^{(n-1)}_{n-1,n})(G^{(n-2)}_{n-2,n-1}) \ldots (G^{(2)}_{2,3})(G^{(1)}_{1,2} \ldots G^{(1)}_{n-1,n}), \tag{17}$$

hence having only $2n-3$ Givens transformations. The interpretation is the following. First we apply on $A \in \mathcal{M}_{\text{pure}}$ a sequence of Givens transformations $G^{(1)}_{i-1,i}$, $i = n, \ldots, 3$ going from bottom to top of the matrix, and transforming A into a *Hessenberg* matrix (since the Rk 1 structure blocks are now transformed into blocks of zeros, with one row less). Next we apply a sequence of Givens transformations $G^{(j)}_{j,j+1}$, $j = 1, \ldots, n-1$ going from top to bottom. These are intended to transform the Hessenberg matrix into upper triangular form.

When applying all these Givens transformations, two situations may occur: (i) We have that $A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{bmatrix}$ where $A_{1,1}$ is square of size k by k, for certain k. In this situation, we will probably want to apply the QR-algorithm on each of the submatrices $A_{1,1}$ and $A_{2,2}$ since the eigenvalue problem, the ultimate aim of the QR-algorithm, can then be reduced to two smaller subproblems. (ii) No such zero block exists below the diagonal of A. Then the left induced pure structure blocks are all of maximal allowed rank equal to 1 and hence Theorem 18.2 implies that the complete structure $\mathcal{R}$ is preserved by following the sparse Givens pattern (17).

However, there is one specific case where Theorem 18.2 will fail, namely when $\lambda_k = 0$. Of course this is not essential since Theorem 18.3 states that in this case, we can just choose an effectively eliminating QR-decomposition of A to preserve structure. This can be realized here by additionally choosing $G^{(j)}_{i-1,i} = I_2$ whenever possible for all pairs (i, j) occurring in (17).

Still we may mention that for this specific case, we do not really need to apply an effectively eliminating QR-decomposition of A. Instead, the following theorem shows that the sparse Givens pattern (17) itself is already enough to guarantee the complete preservation of the structure $\mathcal{R}$!

Theorem 19 *Suppose that $\mathcal{B}_k$ is a pure structure block which intersects the diagonal in precisely one element, i.e. $\lambda_k = 0$ and $i_k = j_k =: d$. Let $A \in \mathcal{M}_{\mathcal{B}_k}$. Suppose that we apply a QR-step without shift on A, satisfying the sparse Givens pattern induced by $\mathcal{B}_{k,a} := (d, d-1, r_k)$ and $\mathcal{B}_{k,b} := (d+1, d, r_k)$, both lying just below the diagonal: see Figure 6. Then Theorem 18.2 will essentially remain valid, i.e. if $\mathcal{B}_{k,a}$ has its maximal allowed rank, then the complete structure block $\mathcal{B}_k$ will be preserved.*

Remark. Note that $\mathcal{B}_{k,a}$ and $\mathcal{B}_{k,b}$ are not really the 'induced pure structure blocks' of $\mathcal{B}_k$, since for $\lambda_k = 0$ we defined only one induced pure structure block, being $\mathcal{B}_k$ itself! The strength of the theorem is that in this case, we can work with the weaker blocks $\mathcal{B}_{k,a}$ and $\mathcal{B}_{k,b}$.

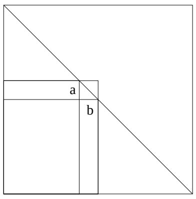


Figure 6: Given the huge pure structure block $\mathcal{B}_k$, the figure indicates the position of the two blocks $\mathcal{B}_{k,a}$ and $\mathcal{B}_{k,b}$.

PROOF. Let $A \in \mathcal{M}_{\mathcal{B}_k}$ be such that $A|_{\mathcal{B}_{k,a}}$ is of maximal allowed rank r_k. Then since

$$r_k = \text{Rank } A|_{\mathcal{B}_{k,a}} \leq \text{Rank } A|_{\mathcal{B}_k} \leq r_k, \tag{18}$$

we must have equality of the two middle ranks, and hence the dth column of $A|_{\mathcal{B}_k}$ is a linear combination of columns $1, \ldots, d-1$.

First we consider the case where $A|_{\mathcal{B}_{k,a}\cap\mathcal{B}_{k,b}}$ is of rank $< r_k$, and hence also $A|_{\mathcal{B}_{k,b}}$ being of rank $< r_k$, by what we just told. It follows then that $A \in \mathcal{M}_{\mathcal{B}_{k,c_k}}$ for *any* value of $c_k \in \mathbb{C}$, where $\mathcal{B}_{k,c_k}$ is defined from $\mathcal{B}_k$ by putting the shift element λ_k equal to c_k. Then Theorem 18.2 guarantees that the new QR-iterate $RQ \in \mathcal{M}_{\mathcal{B}_{k,c_k}}$, at least for every *non-zero* choice of $c_k \in \mathbb{C}$. Clearly, by continuity the same must hold for the value $c_k = 0$, hence proving the theorem.

Now we prove the more difficult case where

$$\text{Rank } A|_{\mathcal{B}_{k,a}\cap\mathcal{B}_{k,b}} = r_k. \tag{19}$$

We want to take over the proof of Theorem 18.2: we define the family R_ϵ as usual, by replacing each zero diagonal element of R with index in $\mathcal{I}_{\text{left},k}$ by the parameter ϵ. But now since $\mathcal{B}_k$ is pure, by definition the left index set $\mathcal{I}_{\text{left},k}$ must contain also the index d. Defining $A_\epsilon := QR_\epsilon$ as usual, we can then *not* hope for condition F2 in the proof of Theorem 18.2 to be valid anymore. Nevertheless, we claim that

$$A_\epsilon \in \mathcal{M}_{\mathcal{B}_{\epsilon,k}}, \text{ where } \mathcal{B}_{\epsilon,k} = (d, d, r_k, \epsilon c) \text{ for certain } c \in \mathbb{C}. \tag{20}$$

Assuming this for the moment, then Theorem 2 ($\mathcal{I}_{\text{left},k} \cap \mathcal{I}_{\text{dep},A_\epsilon} = \emptyset$) implies that also the new QR-iterate of A_ϵ satisfies $\mathcal{B}_{\epsilon,k}$, for all $\epsilon \neq 0$. Taking limits for $\epsilon \to 0$, the new QR-iterate of $A = \lim_{\epsilon\to 0} A_\epsilon$ will satisfy the limiting structure block $\lim_{\epsilon\to 0} \mathcal{B}_{\epsilon,k} = (d, d, r_k, 0) = \mathcal{B}_k$.

Thus we would be finished if we can prove our claim (20). From the proof of Theorem 18.2, it follows that each of the column spaces of $A_\epsilon|_{\mathcal{B}_{k,b}}$, $A_\epsilon|_{\mathcal{B}_{k,a}}$ and (hence) $A_\epsilon|_{\mathcal{B}_{k,a}\cap\mathcal{B}_{k,b}}$ must be invariant from the value of $\epsilon \in \mathbb{C}$. Now we split the dth column of $A_\epsilon|_{\mathcal{B}_k}$ as a (symbolic) sum $\vec{x} + \epsilon\vec{y}$ with $\vec{x}, \vec{y} \in \mathbb{C}^{n-d+1}$. By our assumption on the maximality of rank of $A|_{\mathcal{B}_{k,a}}$, the first term $\vec{x}$ must be a linear combination of the columns of $A|_{\mathcal{B}_{k,a}}$, and hence of $A_\epsilon|_{\mathcal{B}_{k,a}}$ for each $\epsilon \in \mathbb{C}$. So we can forget about this term $\vec{x}$.

Now we treat the second term $\vec{y}$. Since by (19),

$$r_k = \text{Rank } A|_{\mathcal{B}_{k,a}\cap\mathcal{B}_{k,b}} \leq \text{Rank } A|_{\mathcal{B}_{k,b}} \leq r_k, \tag{21}$$

we must have equality of the two middle ranks, and thus when dropping its top element, the vector $\vec{y}$ will be a $\mathbb{C}$-linear combination of the columns of $A|_{\mathcal{B}_{k,a}\cap\mathcal{B}_{k,b}}$. Hence a unique correction number $c \in \mathbb{C}$ exists which has to be added to the top element of $\vec{y}$ to extend this $\mathbb{C}$-linear combination to the complete vector $\vec{y}$. Since we had written the dth column of $A|_{\mathcal{B}_k}$ as $\vec{x} + \epsilon\vec{y}$, and since the column space of $A_\epsilon|_{\mathcal{B}_{k,a}}$ is independent of the value of $\epsilon \in \mathbb{C}$, we can conclude that the shift element of A_ϵ will be precisely ϵc, hence proving our claim (20). □

A counterexample. Finally we give a counterexample of Theorem 18.3 if we skip the condition that the rank structure $\mathcal{R}$ does not imply singularity (Remark 9). Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & \times \\ 1 & 1 & \times \\ 1 & 1 & \times \end{bmatrix} \in \mathcal{M}_{\text{pure}},$$

where $\mathcal{R}_{\text{pure}}$ is defined by the complete first two columns being of rank at most 1. Then we can find an effectively eliminating QR-decomposition of A by solving $A = QR$ with $Q^H = G^{(1)}_{1,2}G^{(1)}_{2,3}$, where $G^{(1)}_{2,3}$ eliminates the $(3, 1)$ and $(3, 2)$ elements and $G^{(1)}_{1,2}$ eliminates the $(2, 1)$ and $(2, 2)$ elements. But now the reader can easily check that in general, the new QR-iterate $RQ = Q^H AQ \notin \mathcal{M}_{\text{pure}}$!

This example shows the necessity of the assumption (10) in order for our theorems to be valid.

6 Conclusion

In this paper we proved that even if A is a singular matrix, a new QR-iterate can be constructed having the same rank structure as the matrix A itself. We introduced the concepts of effectively eliminating QR-decompositions and sparse Givens patterns, and we showed that these have a nice behaviour from both the computational and theoretical point of view. In a future publication, we will show that, apart from the (shifted) QR-algorithm, our definition of structure block is also more or less preserved under matrix inversion.

References

[1] D.A. Bini, F. Daddi, and L. Gemignani. On the shifted QR-iteration applied to Frobenius matrices. Preprint, October 2003.

[2] D.A. Bini, L. Gemignani, and V. Pan. QR-like algorithms for generalized semiseparable matrices. Tech. Report 1470, Department of Mathematics, University of Pisa, 2004.

[3] S. Delvaux and M. Van Barel. Structures preserved by the QR-algorithm. Report TW 399, Department of Computer Science, K.U.Leuven, Leuven, Belgium, August 2004.

[4] D. Fasino. Rational Krylov matrices and QR-steps on Hermitian diagonal-plus-semiseparable matrices. To appear in Numer. Linear Algebra Appl., 2004.

[5] L. Gemignani. A unitary Hessenberg QR-based algorithm via semiseparable matrices. url: fibonacci.dm.upipi.it/∼gemignan/papers/unihess.pdf.

[6] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The Johns Hopkins university Press, third edition, 1996.

[7] W.B. Gragg. The QR algorithm for unitary Hessenberg matrices. *Journal of Computational and Applied Mathematics*, 16:1–8, 1986.

[8] R. Vandebril, M. Van Barel, and N. Mastronardi. An implicit QR algorithm for semiseparable matrices to compute the eigendecomposition of symmetric

matrices. Report TW 367, Department of Computer Science, K.U.Leuven, Leuven, Belgium, August 2003.